

With A Little Help From My Friends: Can Social Navigation Inform Interpersonal Privacy Preferences?

Sameer Patil
patil@uci.edu

Xinru Page
xpage@uci.edu

Alfred Kobsa
kobsa@uci.edu

Department of Informatics, University of California, Irvine

ABSTRACT

Recent privacy controversies surrounding social networking sites demonstrate that the mere *availability* of settings is not enough for effective privacy management. We investigated whether the aggregated privacy choices of one's social circle might guide users in making informed privacy decisions. We conducted an experiment in which users specified preferences for six privacy-relevant settings in Instant Messaging. In one condition, users were provided with information indicating the privacy preferences of the majority of their "buddies." Our results suggest that while this information did influence user choices, the effect was secondary to that of the "privacy-sensitivity" of the system feature controlled by the particular setting. Frequency of IM usage was also associated with privacy choices. The experiment data coupled with user comments suggest several usability improvements in interfaces for specifying privacy preferences.

Author Keywords

Privacy, Instant messaging, Social navigation, Awareness

ACM Classification Keywords

H.5.3 Information Interfaces and Presentation: Group and Organizational Interfaces

General Terms

Design, Experimentation, Human Factors

INTRODUCTION AND RELATED WORK

Users of interpersonal awareness and interaction systems often need to reconcile the awareness benefits of the system with associated privacy risks [5]. Systems typically allow users some control over their privacy by letting them customize privacy-affecting features via settings. However, the mere presence of these settings is insufficient if users are not aware of their existence, do not understand how to use them, or do not use them due to interaction burden. Recent controversies over privacy settings in the popular social networking site Facebook and in the recently launched Google

Buzz service highlight the downsides of confusing and burdensome interfaces for managing these preferences [1, 9]. In fact, these incidents prompted congressional scrutiny [1].

Several prior studies have explored privacy preferences of users (e.g., [3, 12, 13]). The goal of those studies, however, was to understand privacy attitudes and behaviors of users in isolation. In contrast, our study explores whether information about the preferences of their *social circle* could help users make informed choices about their own privacy preferences. While some research has considered social navigation [7] to tackle privacy and security, the social cues used were derived from *all* users of the system [6, 8, 2]. Social-psychological research suggests that, as a guide for their own behavior, people are likely to attribute (consciously and/or subconsciously) more importance to the attitudes and behaviors of those in their social circle(s) [14]. Therefore, we are interested in social navigation cues gleaned from those with whom one interacts rather than from users in general. In this paper, we report on an experiment exploring the effects of such cues on the choice of privacy settings.

METHODOLOGY

The experiment required participants to specify privacy preferences. To avoid biasing participants, we did not reveal our privacy focus. Instead, the study was disguised as a usability evaluation of an Instant Messaging (IM) installer. The "pretend" installer looked and behaved like a standard installation program for Windows[®]. It first showed the license agreement and then a screen for the user to log in to his or her IM account to import the contact list. The installer worked with AOL Instant Messenger[®], MSN Messenger[®], and Google Talk[®], which were identified as the most commonly used based on a pre-survey of our target population. As the final step of the (fake) installation, the installer asked the user to specify preferences for the six privacy settings in Table 1, one at a time in the same order. In the treatment condition, users were shown social navigation cues purportedly indicating the choices made by the majority of their contacts. To increase the believability of the cues, the pre-study instructions informed participants that the study was being conducted in collaboration with IM companies. Participants used their real IM accounts during the installation, which likely boosted the credibility of the cues further. All sessions were conducted in a laboratory by the same experimenter.

After completing the installation, participants filled out two post-study questionnaires: one regarding their IM usage and

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to publish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CSCW 2011, March 19–23, 2011, Hangzhou, China.

Copyright 2011 ACM 978-1-4503-0556-3/11/03...\$10.00.

No.	Description	Question	Options
1.	Login Status	When I log in to Instant Messenger, I set my status to:	- Available - Invisible
2.	Adding Me	Who can add you to their contact list?	- Anyone - Only individuals I explicitly authorize
3.	Inactive Time	Would you like others to be able to view the amount of time you have been inactive (i.e., idle or away)?	- Show to all of my contacts - Show only to these contacts (specify): - Do not show anyone
4.	Saved Conversations	Would you like to be notified when your contacts save the conversation?	- Always notify me - Notify me for these contacts (specify): - Do not notify me
5.	Number of Conversations	Would you like others to be able to view how many IM conversations you are engaged in?	- Show to all of my contacts - Show only to these contacts (specify): - Do not show anyone
6.	Usage Statistics	Would you like others to be able to view statistics about your IM activities (e.g., average signed-in time, inactive periods)?	- Allow all my contacts to view - Allow only these contacts to view (specify): - Do not allow anyone to view

Table 1. Privacy settings included in the installer

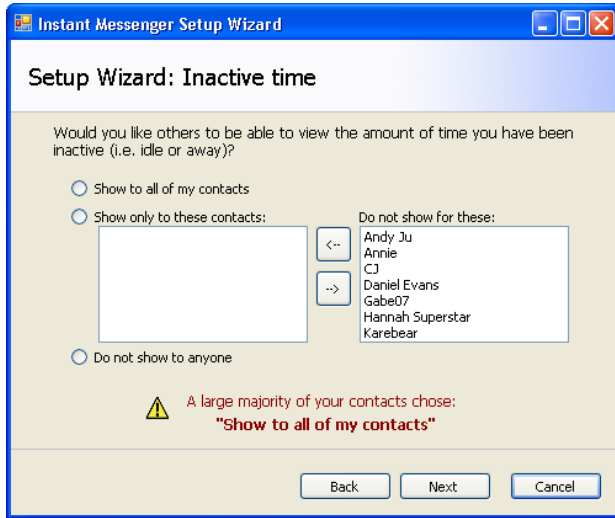


Figure 1. Setting 3 showing a Social Navigation cue in condition T_L

the features encountered in the installer, and another that included the Internet Users' Information Privacy Concerns (IUIPC) scale [11] and the Attention to Social Comparison Information scale [10] as well as demographic questions. We administered the latter questionnaire separately at the end to ensure that its privacy orientation would not bias earlier behavior and responses. Finally, we debriefed each participant to probe for additional comments as well as the credibility of the deception about the purpose of the study. Detailed written debrief notes and the questionnaire responses were analyzed independently by two of the authors.

For each setting, there were two treatment sub-conditions: T_H in which the participants were told that a majority of their buddies chose the option affording higher privacy, and T_L in which they were told that a majority of their buddies chose the lower-privacy option. Figure 1 shows the installer screen for Setting 3 in the T_L sub-condition. Based on believability feedback from pilot tests, we defined three treatment groups

based on the ratio of T_H and T_L settings seen by a participant in the treatment condition. These were: a. 4 T_H & 2 T_L , b. 3 T_H & 3 T_L , and c. 2 T_H & 4 T_L . We used a factorial design to ensure coverage of all combinations of T_H and T_L across the 6 settings in each of the three treatment groups, yielding ${}^6C_4 + {}^6C_3 + {}^6C_2 = 15 + 20 + 15 = 50$ combinations. One participant was assigned to each combination and an additional 18 participants to a control condition without social navigation cues. All assignments were random. 36 participants were male and 32 female, and the average age was 26. Data from 3 other participants in the treatment condition was discarded because their questionnaire responses and/or debrief comments revealed that they had not fallen for the deception regarding the purpose of the study.

Participants were recruited from a large public university and its vicinity using flyers, mailing lists, newsletters, subject pools, and word-of-mouth. We pre-screened potential participants to include only those who were 22 or older and had lived in the U.S. for at least 5 years. This allowed us to control for the effect of age and culture (our past research shows that undergraduates, who are typically younger than 22, tend to exhibit different privacy attitudes than older people). We also ascertained that participants used one of the three supported IM systems (see above) on a (semi-)regular basis.

FINDINGS

First, we performed a chi-squared test to compare the choices made by subjects in the three treatment groups. We found no statistical differences across the treatment groups ($\chi^2 = 1.3$ $p = 0.73$). This indicates that participant choices were *not* influenced by the ratio of high- and low-privacy cues encountered across the six settings. Therefore, we pooled the three treatment versions into a single treatment group ($N = 50$). Further, for each participant, we treated the choice for a single privacy setting as a separate data point. This yielded two sub-conditions for the treatment group: for each setting, half of the participants ($N = 25$) was told that a majority of their buddies chose higher privacy (sub-condition T_H), while the other half was told that a majority chose the lower privacy option (sub-condition T_L).

Next, we examined the choices made by all 68 participants (50 treatment + 18 control). For Settings 3 through 6, participants were given the option to differentiate by individuals (see Table 1). However, this option was chosen for only 3 out of the $68 \times 4 = 272$ total choices (in the control condition for Setting 5, and in T_H and T_L for Setting 3). Moreover, in all three cases, participants specified a single individual, suggesting that higher privacy was desired from all others. Therefore, we recoded these three choices to be the same as the desire for high privacy. This recoding resulted in binary choices (high or low privacy) for all participants for all settings. As a result, we could use binomial instead of multinomial analysis, leading to greater clarity and robustness.

Since a Chi-square test cannot be applied to continuous explanatory variables, we employed regression analysis which allowed us to control for these factors in a single test [4]. We employed binomial logistic regression since it yields less

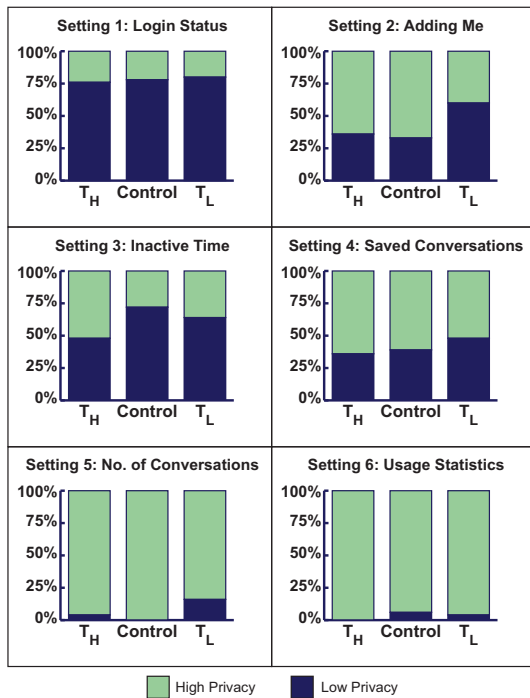


Figure 2. Choices for each setting across the experiment conditions

biased results than ordinary least-squares regression when dependent variables involve binary choices. Further, we incorporated mixed effects, which allowed us to account for repeated measures of the within-subjects design by treating “participant ID” as a random effect. In addition to behavioral data, the regression also included the post-study questionnaire responses. We also considered interaction between these factors, and with treatment sub-conditions and settings. We compared the two models (with and without interactions) with an analysis of deviance using a Chi-squared based estimate. The interaction effects model was not significantly different than one without interactions ($\chi^2 = 7.2$, $p = 0.12$). Thus, the discussion that follows is based on the results of the binomial mixed effects regression without interactions.

A priori, we expected that the choices for each setting in the control condition would be split roughly equally among the high and low privacy options. We further expected social navigation to push the choices toward higher privacy T_H and lower privacy in T_L . Figure 2 shows the results for each setting for the control and treatment conditions. Without social cues, the ratio of those choosing the higher privacy option vs. the lower privacy one fell anywhere between 0 to 1 and was not necessarily near 0.5 as we had anticipated. Further, only settings 1 and 4 show the anticipated trend of social navigation nudging the privacy choices higher or lower, but the magnitude of the effect is small. In fact, a binomial mixed effects logistic regression found no statistically significant differences between the control condition and either treatment sub-condition (T_H vs. C: $p = 0.34$, T_L vs. C: $p = 0.08$). This suggests that privacy choices made without social navigation cues (i.e., in the control condition) may not be directly comparable to those made with the aid of social navigation.

Nonetheless, the differences between the two sub-conditions are statistically significant ($p < 0.0272$): a larger proportion chose the higher privacy option in T_H and vice versa. The effect of social navigation cues, however, is not large. The unexponentiated regression coefficient for sub-condition is 0.56 ($p < 0.05$) reflecting roughly a 75% increase ($(e^{0.56} - 1) * 100\%$) in the odds¹ of choosing the higher privacy option in T_H than in T_L . Instead, the primary driver appears to be the privacy aspect controlled by the setting. Figure 2 shows that the proportions of high and low privacy choices are not the same across the settings, thereby indicating that the underlying features differ in “privacy sensitivity” ($p < 0.001$ except Setting 3 with $p < 0.014$). Repeated t-tests largely yielded the same result after Bonferroni correction ($p = 0.05/15 = 0.003$). At the same acceptance level, the distributions of Settings 1 & 3 ($t = -2.3$, $p = 0.03$), 2 & 4 ($t = -0.34$, $p = 0.73$), and 5 & 6 ($t = -1.16$, $p = 0.25$) are statistically indiscernible. Settings 1 & 3 appear to be less privacy sensitive (intercept, 1.03), 5 & 6 highly sensitive (4.48, 5.44), and 2 & 4 (1.83, 1.97) in the middle. Table 1 reveals that these groups of settings indeed control similar aspects: Settings 1 & 3 refer to the visibility of status, Settings 2 & 4 control awareness of privacy-affecting actions of one’s contacts, and Settings 5 & 6 deal with the analysis of IM activities. We did not find effects of a participant’s total IUICP score [11] and Attention to Social Comparison Information score [10]. However, login frequency was negatively associated with selected privacy levels, i.e., those who logged in daily chose lower privacy compared with those who logged in less frequently (coefficient = -0.47 , $p < 0.0201$). This suggests that infrequent IM users are more privacy-sensitive in a way not measured by the IUICP, or that different levels of use lead to different privacy desires.

Open-ended questionnaire responses and post-study debriefings also provided useful insights. Several participants noted the benefits of adjusting privacy preferences during installation. Their comments indicated that it would raise awareness of potential privacy risks and the existence of corresponding settings. They anticipated refining their preferences upon more experience in using the system. Further, numerous remarks on the usability of the installer corroborated the effectiveness of our deception about the purpose of the study.

IMPLICATIONS

Our findings show that social navigation cues derived from privacy attitudes and practices of one’s social circle can serve as useful *additional* guidance when specifying one’s own privacy preferences. Their influence, however, is secondary to that of the privacy-sensitivity of the feature itself. The utility derived from the cues may be further limited for settings that show floor or ceiling effects (see e.g., Settings 5 and 6 in Figure 2). This suggests that primary importance ought to be placed on the privacy-sensitivity of a setting.

¹A 75% increase in the *odds* differs from a 75% increase in actual choices. Translating odds to predicted probabilities shows that the effect is not large; e.g., if the probability of a person choosing the high privacy option in T_L is 0.4, then this will increase to 0.54 in T_H .

Additionally, 5 of the participants served a note of caution that social navigation could confuse or annoy some users and be perceived as patronizing. For example, one of them commented: “That [social navigation] is not needed. These are my preferences, not theirs.” Another shared: “I felt uncomfortable with it [social navigation]. Remove and let me make my own decisions.” Other comments criticizing the look & feel of the cues as too intrusive indicate the need to make the cues subtle yet easy to notice and understand.

The data point to the possibility of grouping settings with similar privacy-sensitivity levels. This could help deal with the large number of privacy settings offered by current systems like Facebook. The groupings could then be assigned similar defaults and policies. For instance, defaults for highly privacy-sensitive features may follow an opt-in policy, while an opt-out approach might be preferable for less privacy-sensitive aspects. Our results also point to the utility of asking for privacy preferences at setup². If the number of privacy settings is too high then only the most privacy-critical settings could be included into the setup.

The option for customizing preferences for individual contacts was hardly used. As uncovered by prior research [13] and echoed by some participants, customization at the group level may be sufficient³: “Usually people aren’t going to want to take the time to go through their friends one by one. So maybe it could have the option to do it by groups as well.” However, several participants preferred a universal setting with specific exceptions (e.g., boyfriend, stalker, etc.): “It’s an individual that matters, not a group.” This is reflected in the three contact-level choices; these allowed access to a single buddy while maintaining privacy from all others. This points to the utility of a “whitelist” that is the reverse of the standard “block” feature. This would allow “full access” to certain individuals (or groups) with a single click.

LIMITATIONS

Extension of our findings to populations other than U.S. adults requires further empirical validation. Further, a larger sample size could enable a deeper exploration of data subsets (e.g., an independent analysis of choices for each setting).

CONCLUSION

We explored the utility of social navigation for interpersonal privacy management. We found that aggregate information about the privacy choices of one’s social circle can be a secondary source of guidance, provided the cues are non-intrusive and easily understandable. Of primary importance, however, is the privacy-sensitivity of the system feature. Seeking preferences for privacy settings during setup could raise user awareness of associated issues and enable more informed privacy management without undue burden. Although the results were derived from IM, similar settings are present in most interpersonal awareness and interaction

²A lack of such transparency was a major contributor to user confusion and backlash against Facebook and Google Buzz.

³Unfortunately, due to the nature of the installer and the associated IM protocols, we did not have access to information about whether and how the participants organized their contacts into groups.

systems. Therefore, we believe that the findings will apply broadly to all systems with awareness-privacy tensions.

ACKNOWLEDGEMENTS

We thank John Sören Petterson for help in study design, Tijana Gonja and Yen-Sheng Chiang for guidance in analysis, and the anonymous reviewers for thoughtful comments. We also thank the study participants. This research has been supported by NSF Grant Nos. 0205724 and 0808783.

REFERENCES

1. C. Albanesius. Schumer Asks FTC to Investigate Privacy of Facebook, Other Sites. *PC Magazine*, April 2010.
2. A. Besmer, J. Watson, and H. R. Lipford. The Impact of Social Navigation on Privacy Policy Configuration. In *SOUPS '10*, pages 7:1–7:10, 2010.
3. S. Consolvo, I. E. Smith, T. Matthews, A. LaMarca, J. Tabert, and P. Powledge. Location Disclosure to Social Relations: Why, When, & What People Want To Share. In *CHI '05*, pages 81–90, 2005.
4. M. J. Crawley. *The R Book*. John Wiley and Sons, illustrated, reprint edition, 2007.
5. M. J. Culnan and P. K. Armstrong. Information Privacy Concerns, Procedural Fairness, and Impersonal Trust: An Empirical Investigation. *Organizational Science*, 10(1):104–115, 1999.
6. P. DiGioia and P. Dourish. Social Navigation as a Model for Usable Security. In *SOUPS '05*, pages 101–108, 2005.
7. P. Dourish and M. Chalmers. Running Out of Space: Models of Information Navigation. In *Short paper presented at HCI '94*, 1994.
8. J. Goecks, W. K. Edwards, and E. D. Mynatt. Challenges in Supporting End-user Privacy and Security Management with Social Navigation. In *SOUPS '09*, pages 1–12, 2009.
9. M. Helft. Critics Say Google Invades Privacy With New Service. *The New York Times*, February 2010.
10. R. D. Lennox and R. N. Wolfe. Revision of the Self-Monitoring Scale. *Journal of Personality and Social Psychology*, 46(6):1349–1364, 1984.
11. N. K. Malhotra, S. S. Kim, and J. Agarwal. Internet Users’ Information Privacy Concerns (IUIPC): The Construct, the Scale, and a Causal Model. *Information Systems Research*, 15(4):336–355, December 2004.
12. J. S. Olson, J. Grudin, and E. Horvitz. A Study of Preferences for Sharing and Privacy. In *CHI '05*, pages 1985–1988, 2005.
13. S. Patil and J. Lai. Who Gets to Know What When: Configuring Privacy Permissions in an Awareness Application. In *CHI '05*, pages 101–110, 2005.
14. J. W. Thibaut and H. H. Kelley. *The Social Psychology of Groups*. Transaction Publishers, 1986.